REVIEW ARTICLE

Sixty years of second language aptitude research: A systematic quantitative literature review

James Chalmers[1] | **Susana A. Eisenchlas**[2] |
Andrew Munro[2] | **Andrea C. Schalley**[3]

[1]Library and Learning Services, Torrens University, Fortitude Valley, Queensland, Australia

[2]School of Humanities, Languages and Social Science, Griffith University, Nathan, Queensland, Australia

[3]Department of Language, Literature and Intercultural Studies, Karlstad University, Karlstad, Sweden

Correspondence
Andrea C. Schalley, Department of Language, Literature and Intercultural Studies, Karlstad University, 651 88 Karlstad, Sweden.
Email: andrea.schalley@kau.se

Abstract

Second language (L2) aptitude has been broadly defined as the rate and ease of initially acquiring a second language. Historically, L2 aptitude has been understood as a stable trait that predetermined L2 achievement, regardless of individual learners' efforts to acquire an L2. This traditional view of L2 aptitude as fixed and stable has led to it being a relatively neglected area of research within second language acquisition (SLA) studies. The little research that was in fact conducted was diagnostic in nature, and mostly used tests such as the Modern Language Aptitude Test (MLAT) to select potentially gifted L2 learners. Given that six decades have passed since the publication of the MLAT, now is a good time to revisit the literature and investigate whether L2 aptitude continues to be viewed as an individual difference of little interest to SLA research. While summative literature reviews of L2 aptitude research have been written, few systematic reviews exist. This article conducts a systematic quantitative literature review (SQLR) to provide a principled, comprehensive and reproducible synthesis of research into L2 aptitude published over the last 60 years (1959–2019). In this SQLR, close to one

hundred journal articles and PhD dissertations were examined to discern generalisations and limitations in the field. This SQLR identifies a shift in the rationale for L2 aptitude testing, in which a diagnostic focus has been replaced by an explanatory perspective. Furthermore, our article points to a renewed interest in L2 aptitude research, which has come to be characterised by a more nuanced and sophisticated understanding of the concept and its components.

1 | INTRODUCTION

Second language aptitude (henceforth, L2 aptitude) has been broadly defined as the rate and ease of initially acquiring a second language (Carroll, 1981). Paradoxically, while many scholars consider aptitude as one of the main predictive individual variables in second language acquisition (SLA), there has been less interest in this variable compared with others (e.g., motivation; see Dörnyei, 2010) as shown by the dearth of empirical studies published until recently (Wen et al., 2017). This situation may stem from the assumption that aptitude is an immutable 'talent' or 'gift' (Mitchell et al., 2019) that learners either have or lack, with little room for enhancement. A second factor explaining the limited interest in aptitude research has been its perceived link with Audiolingualism (Skehan, 1998), leading scholars to question its relevance to current language teaching methodologies.

While new perspectives have emerged since Carroll and Sapon (1959) published the Modern Language Aptitude Test (MLAT), most narrative style reviews of L2 aptitude (e.g., Spolsky, 1995; Wen et al., 2017) still focus on traditional questions, such as whether aptitude predicts language learning success, whether it is modular or monolithic in nature, or whether it is amenable to experience. But new questions are now emerging, setting new research agendas (cf., Dörnyei, 2010; Robinson, 2012; Sawyer & Ranta, 2001; Skehan, 1998, 2002; Wen et al., 2019). In terms of systematic reviews of the field, three meta-analyses have been conducted (Li, 2015, 2016, 2017), which make valuable contributions but are limited to correlational data and thus fail to give a comprehensive overview of the field. In addition, Li (2019) provides a critical overview of the field, combining the results of the meta-analyses with a narrative review.

The 60th anniversary of the publication of the MLAT is thus a timely occasion to revisit the literature on L2 aptitude and to both interrogate the claim that research into second language aptitude has been neglected and investigate whether the motivations for conducting aptitude testing have shifted over the years. These analyses lead us to identify current trends and emerging questions in L2 aptitude research.

To address these issues, we conducted a systematic quantitative literature review (SQLR, Pickering & Byrne, 2014) of empirical research on L2 aptitude. In this systematic review, we establish what topics and issues have been researched, who has undertaken this research, where it was published, at which geographical locations the studies have been conducted, and the types of measures and methods used. This allowed us to identify emerging trends and future directions for research. This article thus makes a critical contribution to the field by mapping the current status of the literature on L2 aptitude. It will be of relevance to researchers new to this field, who

can quickly gain an effective understanding of some of the main research questions, and to researchers already in the field, who may find new insights to inform their own work.

The article is organised as follows. First, we introduce the concept of L2 aptitude. This is followed by a discussion of the methodology of designing and conducting a systematic review to ensure its objectivity and reproducibility. We then present the results of the SQLR, summarising the general characteristics of L2 aptitude research, and discuss the findings of the review in the context of current L2 aptitude research. We argue that L2 aptitude research is not 'dead' but has continued over the past 60 years, albeit scantly at times. Importantly, the data also show that the understanding of aptitude has become more nuanced and thus deserves further study.

1.1 | Second language aptitude

L2 aptitude is an umbrella term that broadly refers to a 'talent' for language learning. In terms of L2 aptitude research, this talent is defined as the ease and rate at which the L2 is initially acquired (Carroll, 1981). Based on this definition, the term 'aptitude' appears to refer to a real-world phenomenon, but fails to describe it in any objective or meaningful way. Definitions of L2 aptitude in the literature (cf. Dörnyei, 2010; Robinson, 2012; Sawyer & Ranta, 2001; Skehan, 1998, 2002) suffer from a lack of clarity, with no conceptual consensus emerging (Rogers et al., 2017). Most definitions of L2 aptitude are based on Carroll's work with the MLAT (Carroll & Sapon, 1959) or other instruments that seek to measure this 'talent' for language learning (see Granena, 2013, for the LLAMA test; Grigorenko et al., 2000, for CANAL-FT[1]; and Linck et al., 2013, for Hi-LAB[2]). Indeed, the only point of agreement among researchers appears to be that L2 aptitude is what L2 aptitude tests measure (Dörnyei, 2010; Singleton, 2017). In other words, the theories of L2 aptitude are derived from the power of these tests to predict L2 achievement. As a result, scholars agree that further research is needed to develop a better understanding of the concept of L2 aptitude (Singleton, 2017; Wen et al., 2017).

Much of L2 aptitude research has focused on measurement as a way of screening learners to determine their (un)suitability for foreign language learning instruction (Spolsky, 1995). Yet, while L2 aptitude may not be readily defined, tests such as the MLAT have reliably predicted L2 achievement for over 60 years (Sasaki, 2012). Indeed, L2 aptitude has been identified as the single best predictor of L2 achievement, along with L2 motivation (Dörnyei, 2010; Wen et al., 2017). The MLAT is arguably the benchmark test of L2 aptitude, even after 60 years (Sasaki, 2012). Its benchmark status stems primarily from its high and consistent levels of validity ($r = 0.4$–0.6; Sasaki, 2012) and reliability ($r = 0.55$–0.92; Sasaki, 2012) when applied to methods of instruction beyond Audiolingualism, as well as to both formal and naturalistic contexts of acquisition (Sawyer & Ranta, 2001).

Once the MLAT was fully developed and standardised, Carroll (1962) reverse-engineered a post-hoc formulation of L2 aptitude through factor analysis of test data to delimit his construct of L2 aptitude to a subset of cognitive and perceptual linguistic abilities that lead to faster and easier language learning. This analysis resulted in his four-factor construct of L2 aptitude (see Table 1), which continues to be extremely influential to this day (Skehan, 2002; Wen et al., 2017).

The factor analyses of L2 aptitude measures clearly showed the construct to be componential and not unitary (Carroll, 1981), although early research tended to utilise aggregated L2 aptitude test scores. Further research based on the components of L2 aptitude has also shown that L2 aptitude can be broadly categorised into phonological abilities, language analytic abilities

TABLE 1 Carroll's (1962, pp. 129–130) four factors of L2 aptitude

Factor	Description
Phonetic coding	*the ability to meaningfully store auditory information for access at a later time*
Grammatical sensitivity	*the ability to "handle grammar" and discern the functions of words in various contexts*
Rote memory for foreign language materials	*the ability and capacity to memorise a number of associations from the input materials, similar but independent to phonetic coding, encompassing more than just phonetic information*
Inductive language learning	*the ability to infer linguistic forms, rules, and patterns from new linguistic content [...] with a minimum of supervision or guidance*

(comprising inductive language learning and grammatical sensitivity) and memory abilities (Sasaki, 1993a, 1993b; Skehan, 2002).

In summary, L2 aptitude is a complex construct that subsumes various abilities implicated in L2 learning. These abilities are most readily assessed through tests that purport to measure the construct, although the MLAT-like tests lack a rigorous theoretical basis. L2 aptitude is thus an umbrella term that covers various sub-components.

2 | METHODOLOGY

A systematic quantitative literature review (SQLR) was performed following the well established guidelines set out in Pickering and Byrne (2014). A SQLR aims to (1) provide a comprehensive mapping of a field of inquiry; (2) provide an explicit and reproducible method for identifying and selecting literature; (3) summarise the field at the 'big picture' level; (4) extend beyond correlational data while offering a quantitative view of the field; and (5) uncover broader generalisations and limitations of the field.

To be included in the current review, each source had to meet all four of the following criteria: (1) be a journal article or a PhD thesis[3]; (2) pertain to empirical research that included some aspect of L2 aptitude; (3) employ one of the established and readily available instruments for measuring L2 aptitude[4]; and (4) be published in English.[5]

Original research papers and theses presenting empirical studies were obtained from the following scholarly electronic databases: Eric, JSTOR, ProQuest, SAGE, Web of Science and Wiley.[6] Searches were periodically conducted across all the databases between May 7, 2016 and August 24, 2020.

For each database, the same keyword search was employed (see Table 2). Keywords from column 1, which identify L2-aptitude-specific measures and tests, were combined with the keywords from column 2, which delimit the field of research to second language learning. The keyword phrases in Table 2 are written exactly as they were used in the keyword search. For example, the first search for each database was constructed by combining keywords from column 1 with keywords from column 2, for example (*"modern language aptitude test" OR "MLAT") AND ("second language learning" OR "second language acquisition" OR "foreign language learning"*).

Based on the guidelines in Pickering and Byrne (2014), the following data categories were created: publication details, geographical information, research methods used, participant demographic information, and variables measured. Data were extracted from each paper according to the categories and definitions detailed in Appendix A.

TABLE 2 Keywords used in database searches

Keywords	1.	2.
	("modern language aptitude test" OR "MLAT")	AND ("second language learning" OR "second language acquisition" OR "foreign language learning")
	("Pimsleur language aptitude battery" OR "PLAB")	
	("defense language aptitude battery" OR "DLAB")	
	("army language aptitude battery" OR "ALAT")	
	("VORD" OR "York language aptitude test")	
	("cognitive ability for novelty in acquisition of language" OR "CANAL-F")	
	("LLAMA" AND test)	
	"lunic language marathon"	

2.1 | Study selection

As Figure 1 shows, from the initial screening of the keyword database searches, 646 unique sources were selected and assessed for eligibility. Of these 646 studies, 76 were excluded for being non-empirical (e.g., theoretical reviews, book reviews), book chapters, or unrelated. Of the remaining 570 sources, 477 were excluded for not utilising an established and readily available test.

In this SQLR, the term 'source' refers to a publication, either a journal article or a thesis, and the term 'study' refers to research where data were collected from a group of participants and used for analysis. A source can report on more than one study, with each study being recorded separately in the review data set, as long as the participants and/or the data differ between the studies. This situation arose where replications or follow-up studies had been conducted. VanPatten et al. (2013), for instance, reported on four related studies, each with a variation in research design, e.g. a different target grammatical construction, and with a different sample of participants. In total, the review covers 93 sources and 111 studies (for a complete list of references included in the review, see Appendix B).

3 | RESULTS

Following Pickering and Byrne's (2014) guidelines, the results of the review are structured as follows. We first consider the publication details, with a focus on journals, types of publication, and authors most prevalent in L2 aptitude research. We then explore the geographical spread of L2 aptitude research, looking at the institutional affiliation of the researchers conducting the studies as well as the sites of data collection. Demographic information about participants in L2 aptitude studies is considered, followed by a discussion of the research methods used in the original studies. Finally, we summarise key trends, challenges and future research directions.
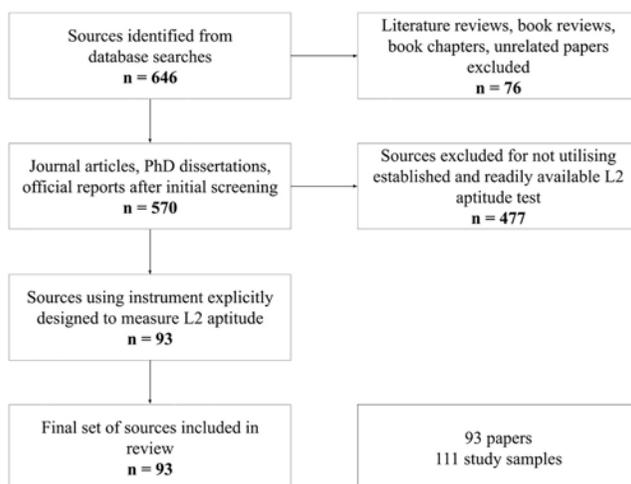
FIGURE 1 Overview of the exclusion/inclusion process

3.1 | Publication details of L2 aptitude research

As evidenced by Table 3, *The Modern Language Journal* and *Language Learning* account for a combined 36.8% of all journal articles published on L2 aptitude research over the period of the review. Overall, the top six journals publishing L2 aptitude research account for 73.7% of all sources published in the review, excluding 17 dissertations. The remaining 15 journals have not published more than one or two studies each.

As Figure 2 shows, *The Modern Language Journal* has also been the most consistent publisher of L2 aptitude research over the period under consideration. From 2013, however, *Language Learning* and *Studies in Second Language Acquisition* were the most active journals publishing research on L2 aptitude.

Figure 3 indicates that the publication of journal articles from the 1960s to 1990 was sporadic, and only three dissertations were published during that period. During the 1990s, however, dissertations became an important source of research in the field, while a low number of articles continued to appear in SLA journals. This situation is changing, with a comparatively sharp increase in journal articles published since 2012 and a relative decrease in the number of dissertations.

TABLE 3 The top journal publishers of L2 aptitude research from 1959 to 2019 (for the full table of all journal publishers, see Supporting Information S1)

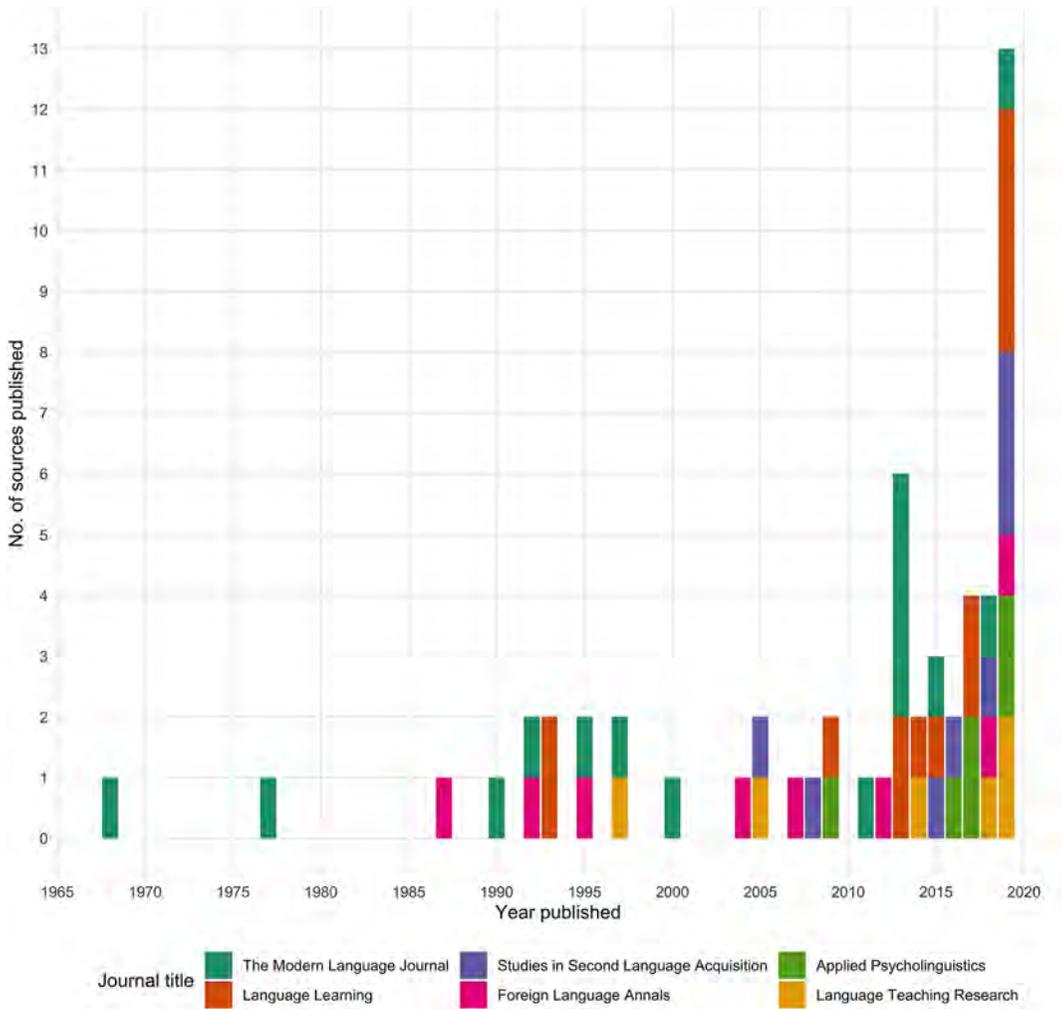| Journal/publisher	Number of studies	Percentage of studies (%)
The Modern Language Journal	15	19.7
Language Learning	13	17.1
Studies in Second Language Acquisition	8	10.5
Foreign Language Annals	8	10.5
Applied Psycholinguistics	6	7.9
Language Teaching Research	6	7.9
Others (*n* = 15)	20	26.3

FIGURE 2 Top journals that published L2 aptitude research, 1959–2019

Table 4 shows that the 14 most published authors account for 36.5% of authorship, while 102 of the 117 remaining authors have only one publication to their name. Note that for the purposes of our SQLR, authorship is individually assigned to each contributing author of a paper. Consequently, Abrahamsson and Hyltenstam (2008), for instance, counts twice: one publication for Abrahamsson and one for Hyltenstam. Eight of the top 14 authors were involved with Sparks and Ganschow's research agenda, which resulted in a number of multi-authored articles (see the references in Appendix B). While this counting may appear to skew the picture, Table 4 clearly evidences that 63.5% of contributions to the field are authored by researchers who have individually published just one or two empirical research papers on L2 aptitude. All of this results in wide variance in authorship that makes it difficult to meaningfully summarise the field.
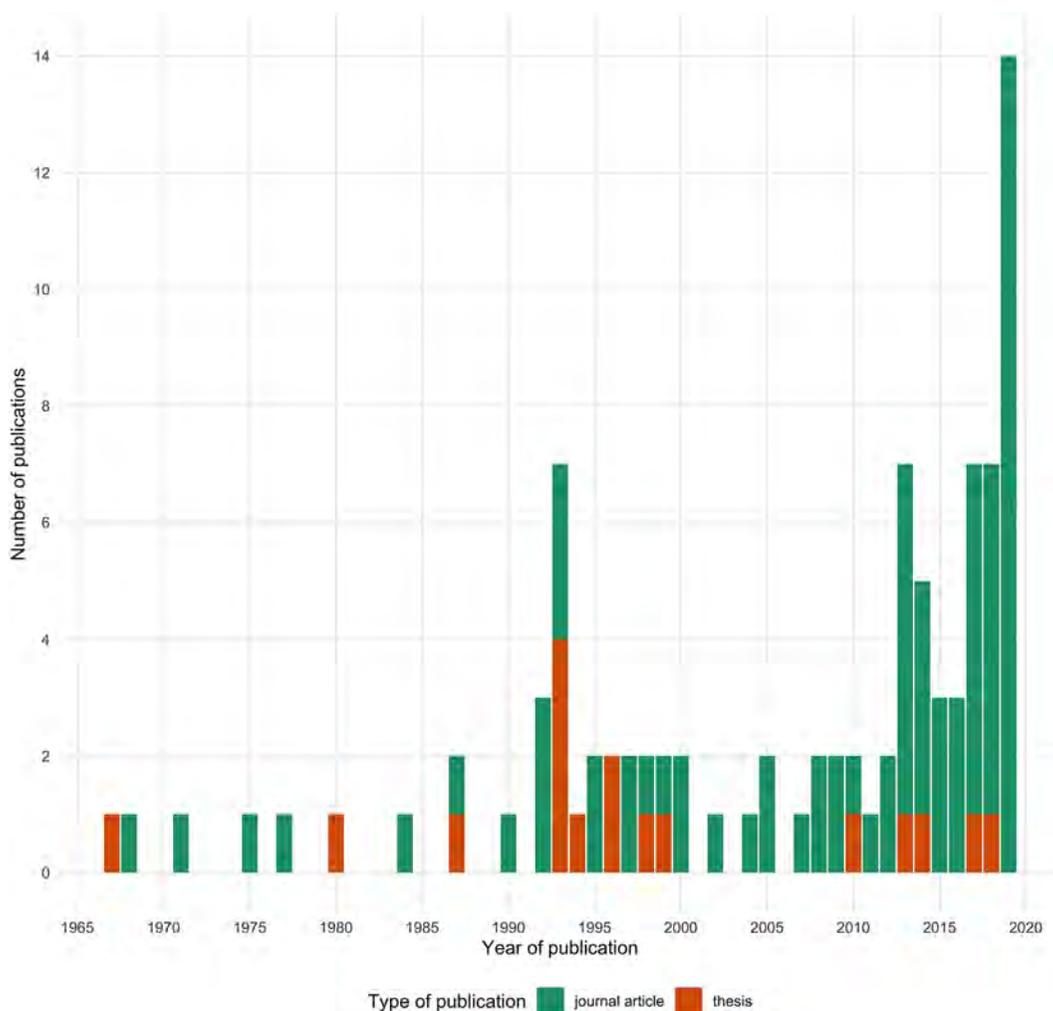
FIGURE 3 Number of publications on L2 aptitude between 1959 and 2019

3.2 | Geographical scope

The geographical scope of L2 aptitude research is summarised in two ways: the count of the countries where an author's research institute was located and a count of the countries where data were collected for each study. Table 5 shows the ordered counts of these two summaries side by side. Clearly, the USA (and North America more broadly) dominate research for both research institutes and data collection sites for L2 aptitude research. Specifically, authors of L2 aptitude research were affiliated with research institutes in the USA 149 times out of a total of 208, or 71.6%. The next most represented country, the United Kingdom, accounted for only 12, or 5.8%. A similar situation is evident for the countries where data were collected. Data for L2 aptitude research were collected from the USA 70 out of 116 times, or 61.9%. The next country from which data were most collected, Japan, accounted for 13, or 11.5%.

Clearly, most of what is known about L2 aptitude comes from the US context. L2 aptitude in the sources reviewed is heavily biased towards language learning in the USA, with the vast majority of

TABLE 4 Most active authors publishing L2 aptitude research 1959–2019 (for the full table of all authors, see Supporting Information S2)

Author	Number of studies	Percentage of studies (%)
R. L. Sparks	17	8.2
L. Ganschow	12	5.8
J. Patton	10	4.8
G. Granena	6	2.9
J. Javorsky	4	1.9
M. Artzer	3	1.4
R. DeKeyser	3	1.4
M. Faretta-Stutenberg	3	1.4
N. Humbach	3	1.4
J. Luebbers	3	1.4
K. Morgan-Short	3	1.4
J. Pohlman	3	1.4
K. Saito	3	1.4
Y. Suzuki	3	1.4
Others ($n = 117$)	133	63.5

TABLE 5 Geographical scope of L2 aptitude research in terms of countries where authors' research institutes were located and countries from which data were collected

Countries of authors' research institutions		Countries from which data were collected	
Country	Total	Country	Total
USA	149	USA	70
UK	12	Japan	13
Japan	9	UK	6
Canada	8	Canada	5
China	7	Turkey	4
Spain	6	Spain	3
Germany	3	China	2
The Netherlands	3	Brazil	1
Fiji	2	Fiji	1
Iran	2	Germany	1
Sweden	2	Greece	1
Australia	1	Iran	1
Ireland	1	Mexico	1
New Zealand	1	The Netherlands	1
Taiwan	1	New Zealand	1
Turkey	1	Sweden	1
		Taiwan	1

research institutions and data collected in the United States of America. Despite the heavy concentration of research published by scholars with US affiliations, Table 5 shows a slightly wider geographical range of countries from which data were collected than the institutional affiliations would suggest.

3.3 | Participants

The SQLR data across studies show significant heterogeneity in the participants' profiles. The average ages of participants across the studies ranged from 6.9 years (Sparks et al., 2009) to 44.1 years (Sheffield, 1993). Participants were only slightly more likely to be female (55.8%) than male (44.2%). While the most-reported educational level of participants was university studies (50.0%), around a third of participants had high school studies only (35.2%), and 0.8% (Dąbrowska, 2018) included participants with no formal education. Aside from English (48.4%), the most common L1s spoken by participants were Japanese (8.2%), Chinese (7.5%) and Spanish (5.7%), with a total of 31 L1s recorded as spoken by participants across all studies (see Supporting Information S3). The most common L2s being learned by participants were Spanish (21.0%), French (15.6%), English (15.1%) and German (9.7%), with a total of 37 L2s recorded as being learned by participants across all studies (see Supporting Information S4).

The level of reporting of participants' demographic information varied significantly across studies (see Table 6). Almost all studies reported on participant numbers (99.1%), participants' education level (93.7%) and L2 being learned (93.7%). The majority of studies reported on participants' gender (70.3%), L1s (67.6%), average age (65.8%), age range (64.0%) and L2 learning experience (58.6%). However, socio-economic status, an important variable in education research (Reardon et al., 2014) was only reported in 16.2% of the studies.

Plainly, the data collected from these participants are not representative of the general population. This is unsurprising, given the context and motivations driving the historical use of L2 aptitude tests. Further research into L2 aptitude could broaden the participant demographics' profiles, as well as improve the reporting practices on demographic information.

3.4 | Research methods

The research methods of L2 aptitude testing reflect its psychometric nature. Typically, studies are quantitative (93.7% of studies) rather than qualitative (1.8%) or mixed (4.5%). The orientation

TABLE 6 Reporting on participants' demographic information

Participant information	Not reported (*n*)	Reported (*n*)	Not reported (%)	Reported (%)
Participant numbers	1	110	0.9	99.1
Education level	7	104	6.3	93.7
L2 being learned	7	104	6.3	93.7
Gender	33	78	29.7	70.3
L1	36	75	32.4	67.6
Average age	38	73	34.2	65.8
Age range	40	71	36.0	64.0
L2 learning experience	46	65	41.4	58.6
Socio-economic status (SES)	93	18	83.8	16.2

towards quantitative studies may be a result of the initial motivation of research in this area, which was diagnostic.

There is a tendency in L2 aptitude studies towards longitudinal (54.1%) rather than cross-sectional (40.5%) or mixed (5.4%) designs. Still, the two orientations are relatively balanced. The categories, however, are not clear cut and much depends on how the term 'longitudinal' is defined (see Appendix A).[7] The most typical studies in this category follow a design in which an initial L2 aptitude test is administered, an educational intervention is conducted, and correlations between aptitude tests scores and final L2 achievement scores are established (Wen et al., 2017).

Although studies in L2 aptitude tend to be observational (58.6%) rather than experimental (41.4%), the number of experimental studies is on the rise, as shown by Figure 4.

Studies of L2 aptitude have differed in their focus, the measures used for determining L2 aptitude, as well as in the dependent and independent variables taken into account. These differences will be discussed in the following subsections. We first consider the foci of L2 aptitude research as well as the measures used, before turning to consider the dependent and independent variables used in L2 aptitude research. Our analysis includes summary statistics and a summary of the trends over time for these areas in L2 aptitude research.

3.4.1 | Foci and measures of L2 aptitude research

In approximately two-thirds of the studies included in the SQLR, L2 aptitude has been used as an explanatory variable for L2 acquisition. As indicated in Table 7, only 30.6% of the studies
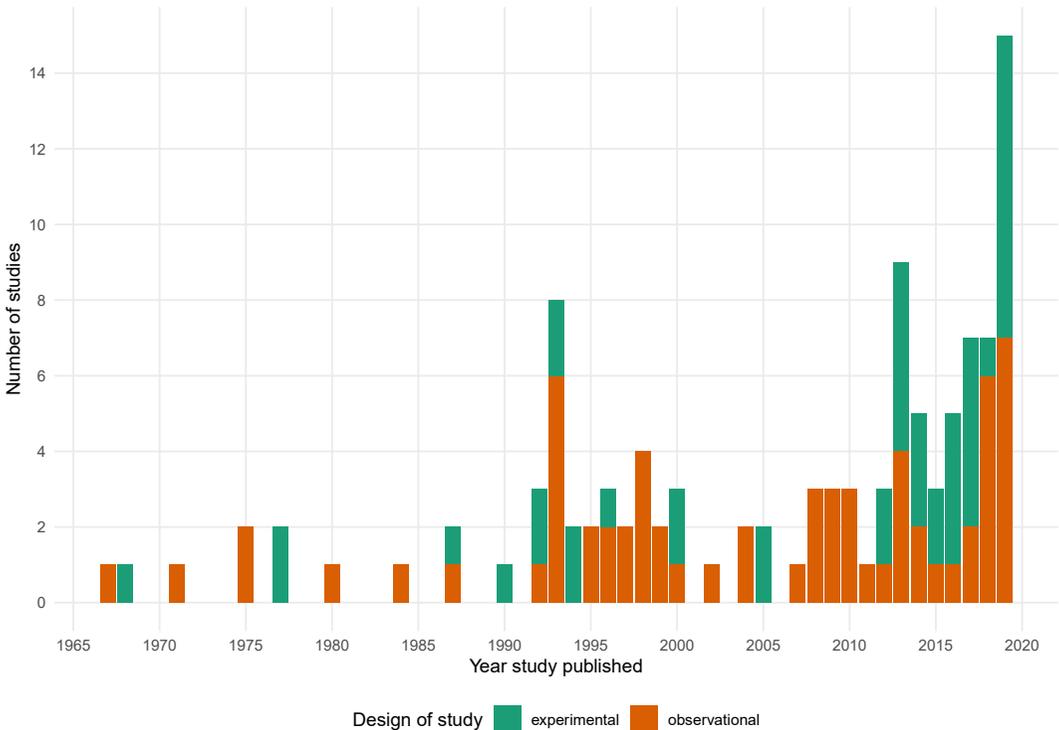

FIGURE 4 Trends in the use of observational versus experimental study designs, 1959–2019

TABLE 7 Focus of L2 aptitude in the research

L2 aptitude focus	Number of studies	Percentage of studies (%)
Explanatory ID variable	70	63.1
L2 aptitude itself	34	30.6
Participant co-variate	7	6.3

have explicitly investigated the nature of L2 aptitude itself and/or of its components (e.g., Cox et al., 2019; Doughty, 2019; Granena, 2016; Li et al., 2019).

Figure 5 shows a shift in focus of L2 aptitude research included in this review. Studies only began to include L2 aptitude as an explanatory variable in the 1990s, with a marked increase in these types of studies occurring over the last decade. In contrast, research into L2 aptitude itself has been occurring since the late 1960s, although the number of studies with this focus has also increased since the 1990s. This trend in the focus of L2 aptitude-related research fits Dörnyei's (2010) observation that the 1960s to the 1990s was a time of research into L2 aptitude test development, with the period from the 1990s onwards, particularly the last decade, being marked by renewed interest in a range of research streams (see Wen et al., 2017). Evidently, research into L2 aptitude has shifted from a diagnostic/predictive focus to one of explanation.

Table 8 shows the usage of the different L2 aptitude measures, with a measure being counted each time it was used for analysis in a study. Some studies analysed the relationship between overall scores from an L2 aptitude battery with overall scores from an L2 achievement test and also analysed individual sub-tests of L2 aptitude with specific aspects of L2 achievement. In such cases, the study was counted twice in Table 8. For example, in Doughty (2019) overall L2 aptitude test scores (MLAT, Hi-LAB) were analysed against successful completion of L2 courses, while sub-tests were analysed as predictors of L2 skills, e.g. speaking and reading. In total, 50 unique measures of L2 aptitude were used in 228 instances (for the full table, see Supporting Information S5).

Table 8 clearly shows two aspects of these data worth noting: the dominance of the MLAT tests as a measure of L2 aptitude, and the use of overall versus specific measures of L2 aptitude. L2 aptitude research has been dominated by the MLAT, which accounts for 58.8% of all L2 aptitude measures included in L2 aptitude research.[8] The LLAMA tests are the second most used measure of L2 aptitude, accounting for 19.7% of all measures of L2 aptitude in the research.

Since 2010, a marked rise in the use of the LLAMA test to determine L2 aptitude is apparent (see Figure 6). Importantly, however, this increase does not co-occur with a decrease in the use of the MLAT tests. Also worth noting is the increased usage over the decade of other measures of L2 aptitude, e.g. CANAL-FT (Grigorenko et al., 2000), PLAB (Pimsleur, 1966) and Hi-LAB (Linck et al., 2013).

The data on the measures of L2 aptitude also show an interesting trend in the use of composite scores – that is, a total score that is the sum of all individual tests in a test battery – and the use of individual test scores. We refer to these types of L2 aptitude measures as 'whole' and 'specific', respectively. Although the MLAT Long, which is a 'whole' measure of L2 aptitude, is the most commonly used measure, the use of specific L2 aptitude measures in research has dominated the field (see Table 9). While many studies included both whole and specific measures of L2 aptitude (e.g., Cummins & Gulutsan, 1975; Doughty, 2019; Sheffield, 1993; Winke, 2013), the inclusion of specific measures suggests that research has overall operationalised L2 aptitude as a componential construct (Wen et al., 2017).
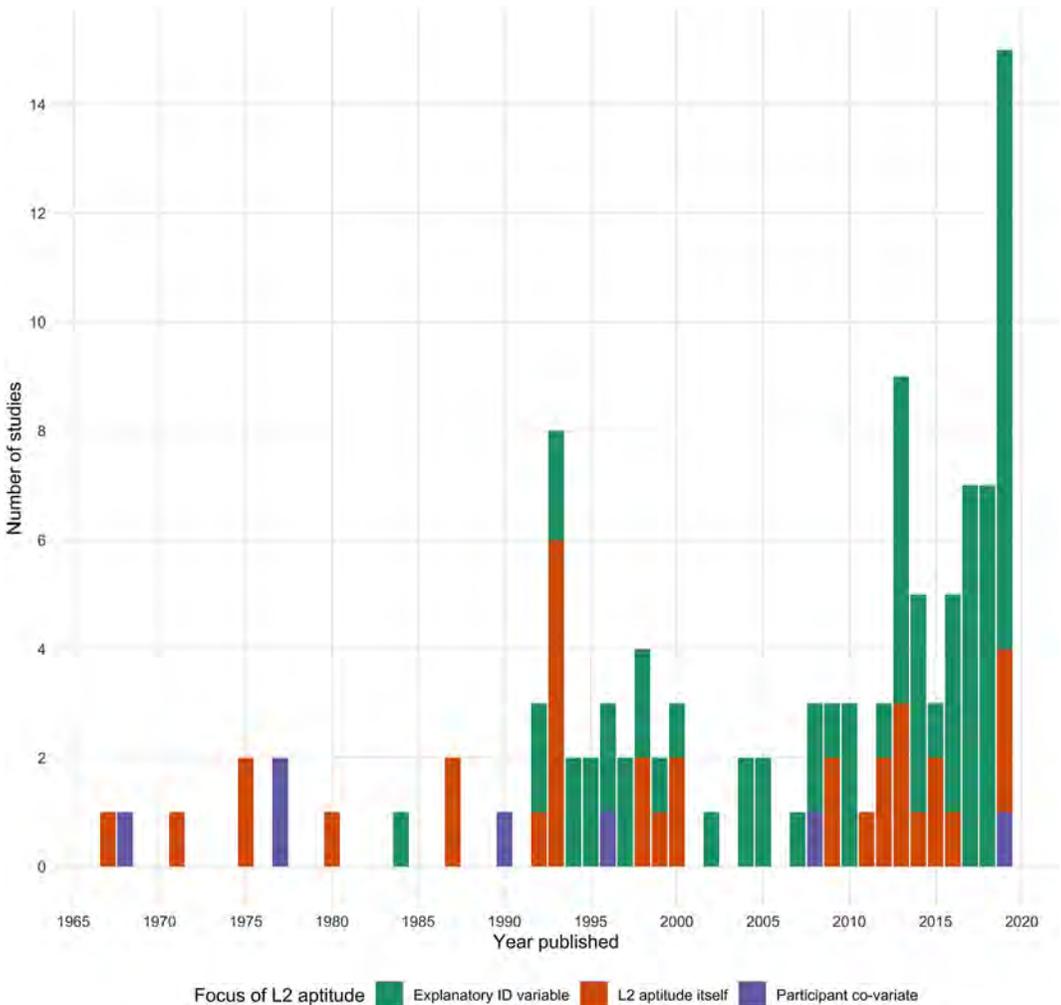
FIGURE 5 Trends in the use of L2 aptitude in the research, 1959–2019

Figure 7 shows the trends in the use of whole compared with specific measures of L2 aptitude. Again, the trend in research over the last decade is towards asking more specific questions on L2 aptitude, with the increasing use of specific measures of L2 aptitude.

3.4.2 | Dependent and independent variables used in L2 aptitude research

Table 10 shows the different types of dependent variables used in the studies included in the review (see Appendix A for definitions for each type of dependent variable). Overall, the four most common dependent variables in L2 aptitude research account for 78.3% of all dependent variables. As expected, all four relate to measuring L2 achievement at a general or specific level, that is, (i) overall L2 achievement (21.7%), e.g. school grades (Muñoz, 2017); (ii) specific L2 skill (20.9%), e.g. L2 listening (Davies, 1971); (iii) specific L2 achievement (18.3%), e.g. L2 pronunciation and fluency (Saito, 2017); and (iv) specific L2 knowledge (17.4%), e.g. explicit L2 metalinguistic knowledge (Granena, 2014). The remaining 21.7% of dependent variables included in L2

TABLE 8 Most prevalent measures of L2 aptitude used in the research

Measure of L2 aptitude	Number of studies	Percentage of studies (%)	Cumulative percentage of studies (%)
MLAT Long	37	16.2	16.2
MLAT IV Words In Sentences	26	11.4	27.6
MLAT V Paired Associates	21	9.2	36.8
MLAT Short	18	7.9	44.7
LLAMA F Grammatical Inference	12	5.3	50.0
LLAMA E Sound-Symbol Association	11	4.8	54.8
LLAMA B Vocabulary Learning	10	4.4	59.2
LLAMA D Sound Discrimination	10	4.4	63.6
MLAT III Spelling Clues	10	4.4	68.0
MLAT I Number Learning	8	3.5	71.5
MLAT II Phonetic Script	8	3.5	75.0
Other ($n = 39$)	57	25.0	100

aptitude research are quite varied and have typically been investigated only once. This focus on L2 achievement is to be expected given that L2 aptitude tests were originally designed to predict L2 achievement (Carroll, 1962).

Figure 8 evidences the trends in the use of the four most common dependent variables in L2 aptitude research for the period under review. It clearly indicates that although overall L2 achievement is still being researched consistently, the last decade has seen a rapid rise in research that includes the more specific types of dependent variables. Once again, the data suggest that rather than merely applying L2 aptitude testing to predict overall L2 achievement scores, L2 aptitude research is becoming more nuanced and focusing on more specific questions.

The six independent variables most frequently included in L2 aptitude research comprise half (50.1%) of all independent variable instances used in the reviewed L2 aptitude research (see Table 11). Another 53 independent variables represent the remaining 49.9% (see Supporting Information S6 for a full overview). A small number of independent variables thus accounts for the majority of research, while numerous independent variables have been considered in very few studies. The six most frequently investigated independent variables comprise variables that were considered in the initial period of test development from 1959 to 1990 (Dörnyei, 2010), and some that were proposed subsequently. Notably, although Carroll (1962) discussed intelligence explicitly and could be said to have discussed affective variables (e.g., attitude, anxiety and motivation) as a function of time spent on task (see Carroll, 1962), L1 skills and working memory were not identified as relevant factors at the time. Figure 9 makes apparent that L1 skills and working memory are more recent developments in L2 aptitude research.
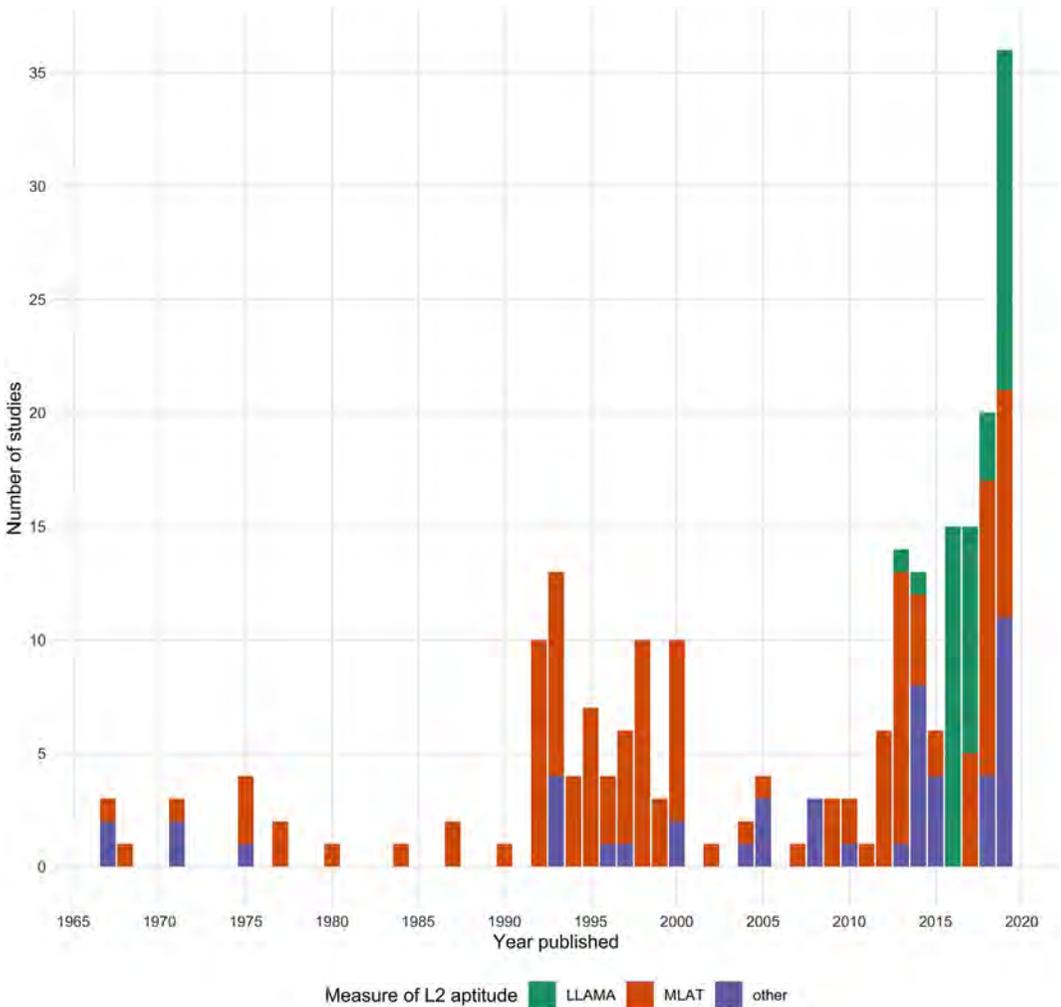
FIGURE 6 Trends in the use of L2 aptitude test measures in L2 aptitude research, 1959–2019

TABLE 9 Use of whole versus specific measures of L2 aptitude across all studies

Type of L2 aptitude measure	Number of instances measure used	Percentage of total instances measure used (%)
specific	154	67.5
whole	74	32.5

Figure 9 shows the trends in the use of the three most common independent variables in L2 aptitude research for the period of this review. Intelligence was the independent variable most consistently included in L2 aptitude research. While L1 skills were most prominent in L2 aptitude research during the 1990s, working memory has become the most researched independent variable of the three over the last decade. Thus, the data support the claim that L2 aptitude research has evolved markedly from Carroll's original four-factor construct (see Carroll, 1962).
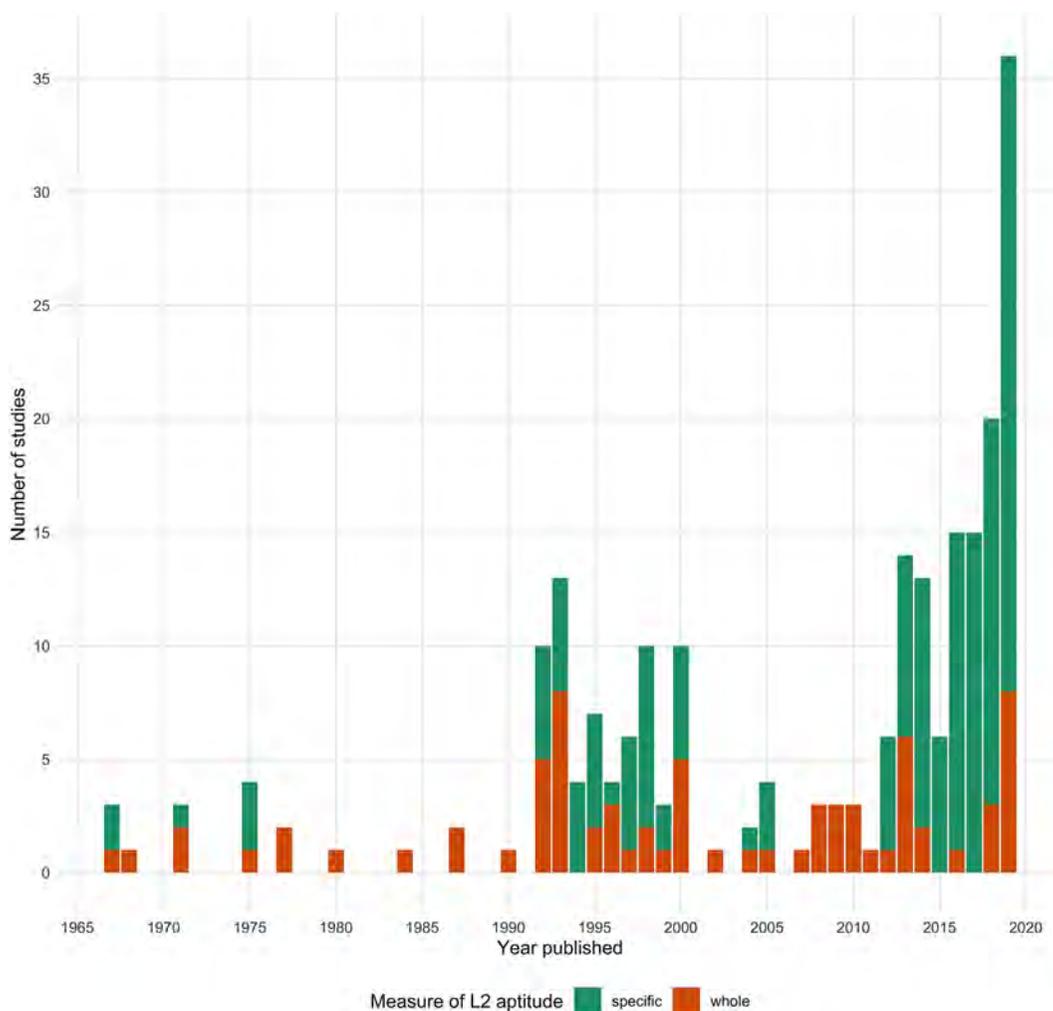
FIGURE 7 Trends in the use of whole versus specific measures of L2 aptitude, 1959–2019

4 | TRENDS, CHALLENGES AND FUTURE DIRECTIONS

The aim of this systematic review was to overview the field of L2 aptitude and document where research has been conducted and disseminated, by whom, on what aspects, by which methods, and what was found. By applying this method, we were able to identify shifts in perspectives and emerging trends, limitations of the findings, and gaps in the literature that merit further investigation. In what follows, we summarise key findings, identify some challenges and suggest areas for future research.

4.1 | Research on L2 aptitude is not 'dead'

The SQLR shows that, despite the limited interest that L2 aptitude has attracted in the SLA literature compared with other individual variables, research has continued to be published, al-

TABLE 10   Dependent variables in L2 aptitude studies

Dependent variable type	Number of studies	Percentage of studies (%)
Overall L2 achievement	25	21.7
Specific L2 skill	24	20.9
Specific L2 achievement	21	18.3
Specific L2 knowledge	20	17.4
L2 aptitude scores	11	9.6
Learning ability/disability	4	3.5
L1 skills	3	2.6
Cognitive styles	1	0.9
Eventual completion of course	1	0.9
L2 experience and development	1	0.9
Neurocognitive processing	1	0.9
On-time completion of course	1	0.9
Set ability	1	0.9
Working memory test scores	1	0.9

beit sporadically, until the 1990s. As seen in Figure 3, that decade saw a resurgence of interest in L2 aptitude research. Interestingly, this resurgence occurred initially in PhD dissertations. More recently, there has been a comparatively sharp increase in the number of journal articles published. Indeed, special issues published since the cut-off date of this review (see Doughty & Mackey, 2021; Li & DeKeyser, 2021) and edited books (e.g., Wen et al., 2019) show the upward trend of research into L2 aptitude continues. This trend may stem from a renewed interest in L2 aptitude, focused on more nuanced questions and novel methodologies, as shown in this overview.

4.2 | Research on L2 aptitude is geographically and demographically limited

From the SQLR, L2 aptitude research can be seen to have a comparatively long, but geographically and demographically limited history. Geographically, L2 aptitude research is heavily biased towards North America, particularly the USA and, to a lesser degree, Canada (see Table 5). Two explanations are possible: (1) that this review was limited to English language journals[9]; and (2) that compared with other countries, there are more academics conducting L2 aptitude research in the USA, typically on data collected from North American participants. Of these participants, the majority is L1 English-speaking high school or university students learning foreign/second languages. Clearly, this data set cannot be considered representative of the wider population of language learners, of which most are not L1 English speakers learning languages at high school or university. Indeed, the question needs to be asked if the published research findings on L2 aptitude are generalisable to diverse populations.
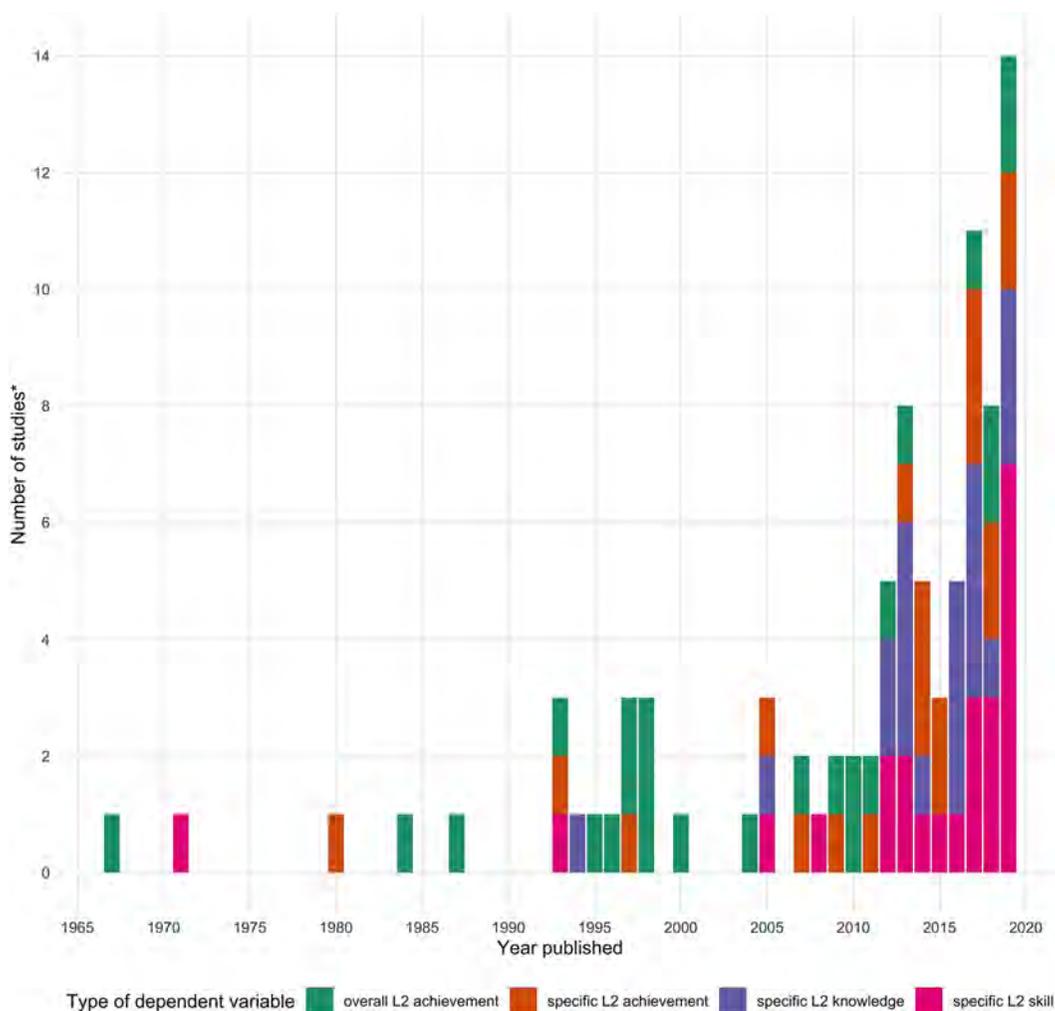
FIGURE 8 Trends in the use of dependent variables in L2 aptitude research, 1959–2019

TABLE 11 The most frequently included independent variables in L2 aptitude studies

Independent variable	Number of studies	Percentage of studies (%)
Intelligence	32	14.5
L1 skills	20	9.0
Working memory	18	8.1
Attitude	15	6.8
Anxiety	14	6.3
Motivation	12	5.4
Age	7	3.2
Gender	7	3.2
L1 vocabulary	6	2.7
Learning style	6	2.7
Others ($n = 49$)	83	38.0

FIGURE 9 Trends in the use of the top three independent variables in L2 aptitude research, 1959–2019

4.3 | Reporting practices are patchy

Participants in this research are poorly understood, as the reporting of their background characteristics is inconsistent. If the background characteristics across samples are unknown, then true comparisons are not possible and thus meta-analyses would be unfeasible based on these data. This challenge, in turn, compromi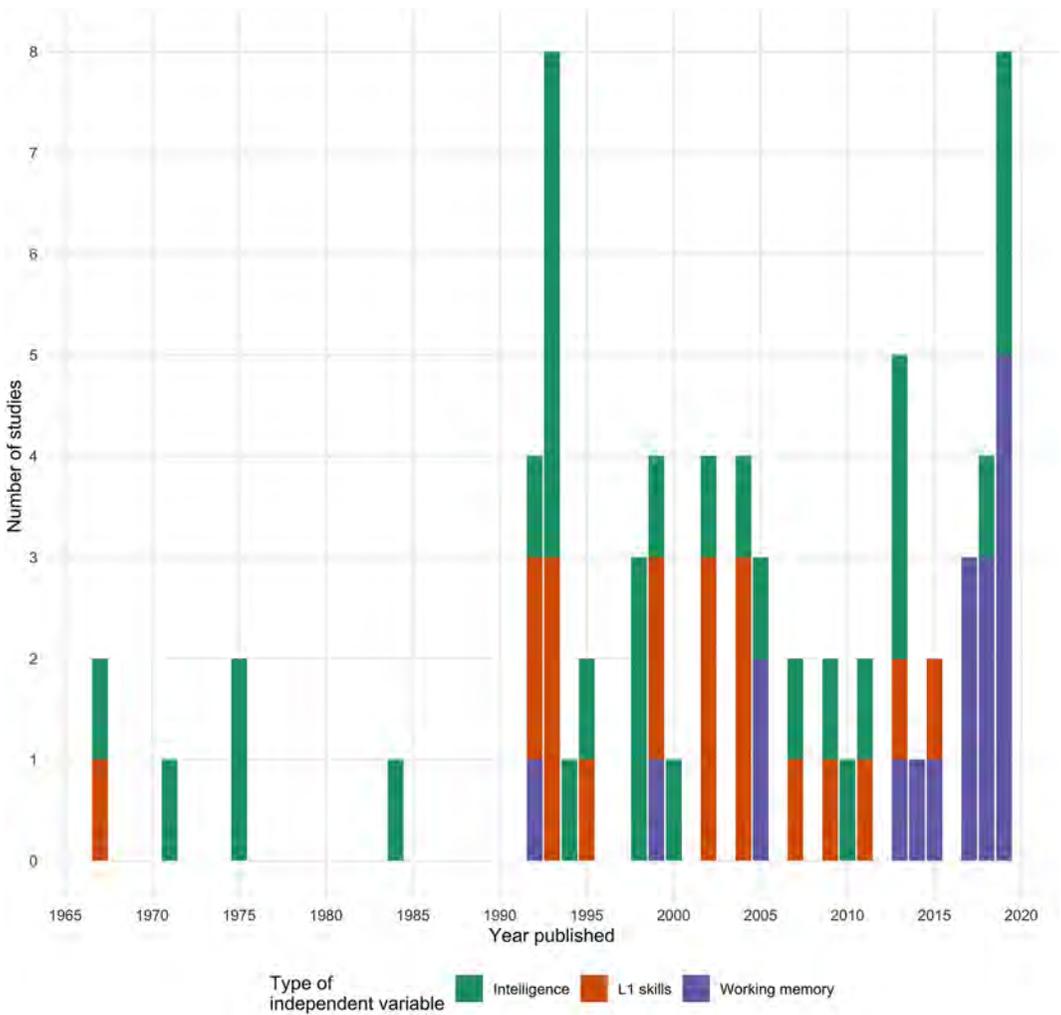ses the interpretability of the results. For example, socioeconomic status is considered an important variable in educational outcomes (Reardon et al., 2014), yet only 16.2% of all studies report on this variable. Additionally, less than 60% of the studies reported on previous language learning experience, despite findings that suggest this experience influences L2 aptitude test scores (Rogers et al., 2017). More research is needed to better understand which background factors are relevant in studying L2 aptitude (Rogers et al., 2016, 2017). Increases in the consistency and details of reporting background characteristics of participants would be especially useful if L2 aptitude research expands beyond the current focus on L1 English-speaking, high-school and university aged students in North America.

4.4 | Interest in L2 research is restricted and fractured

From our SQLR on L2 aptitude research we can see that the field is of primary interest to SLA researchers, as most studies have been published in language learning focused journals (see Figure 2), which is understandable given the original motivation of the field. Moreover, the field of L2 aptitude research appears to be a fractured one, with little consensus or long-term research agendas. Notable exceptions in the period under review are Sparks, Ganschow and their colleagues.

Curiously, the fracturing of the field is further evidenced by the fact that individual researchers who have published PhD theses on L2 aptitude do not appear to have published subsequent research articles in the field. This suggests that most researchers do not focus on L2 aptitude itself, but rather include L2 aptitude as one of the variables of interest in their studies.

4.5 | Research on L2 aptitude is overwhelmingly quantitative

The SQLR makes it clear that, methodologically, L2 aptitude studies are overwhelmingly quantitative, focusing on comparing L2 aptitude scores with other predictors of L2 achievement. The data are thus correlational, examining statistical relationships between aggregated measures of L2 aptitude, L2 achievement, and to some extent intelligence, attitudes and motivation.

Although L2 aptitude research tended to focus on aggregate scores for a battery of tests, more recent research examines specific tests addressing one area of L2 aptitude to explain outcomes in specific areas of L2 learning. Coextensive with this recent research tendency is the increasing use of a range of L2 aptitude tests other than the MLAT. The most well-known of these are the LLAMA tests, which are free and computer-based (Rogers et al., 2017), although a range of other tests also exist. Notably, the development of new L2 aptitude tests is only partially captured in our review. The most prominent of these is the Hi-LAB (Doughty, 2014), which purports to predict high levels of L2 ultimate attainment with tests that measure various aspects of memory (with a particular focus on working memory), implicit learning, processing speed and auditory perceptual acuity (see also Linck et al., 2013). Significantly, all these new measures in L2 aptitude testing continue to be quantitative, reflecting the psychometric origins of L2 aptitude research.

4.6 | Variables investigated

L2 aptitude research has changed its orientation from a predictive to an explanatory focus, which co-occurs with a renewed interest in the field of L2 aptitude research in SLA studies. Variables more recently investigated go beyond Carroll's (1962) original four-factor construct of L2 aptitude and related variables, that is, intelligence and motivation, to include new and important factors such as L1 skills and working memory (see Figure 9). Recent research identifies that a relationship between these variables does exist and is worthwhile investigating further (Sparks & Patton, 2013), given their potential theoretical and practical implications. Theoretically, the focus on new variables should lead us to interrogate precisely what comprises L2 aptitude and to reconsider its constitutive components. This may necessitate the development of new testing instruments. Practically, a reconsideration of the components of L2 aptitude might lead to the identification of strengths and weaknesses in individual learners' profiles, thus enabling a more personalised approach to language instruction.

5 | CONCLUDING REMARKS

The aim of this article was to review 60 years of research into L2 aptitude and to examine whether this construct continues to be viewed as an individual variable of little interest to SLA research. To address this question, we conducted a systematic review of 60 years of empirical research in this field.

The SQLR has shown that, despite popular perceptions, research into L2 aptitude is in fact alive and growing, both in volume and sophistication. Having initially been understood as a four-factor construct, L2 aptitude is currently being interrogated for its potential multifactorial nature.

Arguably, one issue worthy of further research is the relation that may pertain between L2 aptitude and other complex individual variables, such as motivation (Dörnyei, 2010). Another issue that merits investigation is the potential contribution of working memory to L2 aptitude. Indeed, some current research goes as far as to largely conflate working memory with L2 aptitude (Singleton, 2017; Wen, 2016). This in turn raises the intriguing possibility that L2 aptitude could in fact be amenable to training, a prospect which directly contradicts the notion that L2 aptitude is an immutable gift of the fortunate few.

ORCID

James Chalmers https://orcid.org/0000-0002-3233-6129
Susana A. Eisenchlas https://orcid.org/0000-0002-7700-8444
Andrea C. Schalley https://orcid.org/0000-0003-1323-1548

ENDNOTES

[1] The Cognitive Ability for Novelty in Acquisition of Language (Foreign) Test.

[2] The High-Level Language Aptitude Battery.

[3] Chapters were excluded as they are not as consistently available as journal articles and dissertations via online search engines.

[4] This criterion thus excludes tests such as the Hi-LAB which is not readily available as well as measures that are yet to be validated.

[5] We thus acknowledge that our SQLR may not reflect the full scope of work in this field, as studies could have been published in other languages. However, given that English is the international publication language of the field, the restriction to English is unlikely to have substantially skewed the 'big picture' of aptitude research presented herein.

[6] Our selection of databases was informed by advice from university librarians, preliminary searches and careful review of listed references in the articles and theses consulted.

[7] Conducting long term longitudinal studies presents significant challenges. The high attrition rates in language classes make it hard to select and follow participants over lengthy periods of time, but even if this were possible, disentangling aptitude from the effects of instruction seems unfeasible.

[8] This figure includes the MLAT tests, and all tests derived from the MLAT, for example, the MLAT-E and translations.

[9] This limitation, however, would not explain the dearth of studies from other English speaking countries.

REFERENCES

Abrahamsson, N., & Hyltenstam, K. (2008). The robustness of aptitude effects in near-native second language acquisition. *Studies in Second Language Acquisition*, *30*(4), 481–509. https://doi.org/10.1017/S027226310808073X

Carroll, J. B. (1962). The prediction of success in intensive foreign language training. In R. Glaser (Ed.), *Training research and education* (pp. 87–136). University of Pittsburgh Press.

Carroll, J. B. (1981). Twenty-five years of research on foreign language aptitude. In K. C. Diller (Ed.), *Individual differences and universals in language learning aptitude* (pp. 83–118). Newbury House.

Carroll, J. B., & Sapon, S. M. (1959). *Modern language aptitude test*. The Psychological Corporation.

Cox, J. G., Lynch, J. M., Mendes, N., & Zhai, C. (2019). On bilingual aptitude for learning new languages: The roles of linguistic and nonlinguistic individual differences. *Language Learning*, *69*(2), 478–514. https://doi.org/10.1111/lang.12341

Cummins, J., & Gulutsan, M. (1975). Set, objectification and second language learning. *Journal of Psychology*, *10*(2), 91–100. https://doi.org/10.1080/00207597508247323

Dąbrowska, E. (2018). Experience, aptitude, and individual differences in linguistic attainment: A comparison of native and nonnative speakers. *Language Learning*, *69*(S1), 72–100. https://doi.org/10.1111/lang.12323

Davies, A. (1971). Language aptitude in the first year of the U.K. secondary school. *RELC Journal*, *2*(1), 4–19. https://doi.org/10.1177/003368827100200102

Dörnyei, Z. (2010). *The psychology of the language learner: Individual differences in second language acquisition*. Routledge.

Doughty, C. J. (2014). Assessing aptitude. In A. Kunnan (Ed.), *The companion to language assessment* (pp. 25–46). Wiley-Blackwell. https://doi.org/10.1002/9781118411360.wbcla117

Doughty, C. J. (2019). Cognitive language aptitude. *Language Learning*, *69*(S1), 101–126. https://doi.org/10.1111/lang.12322

Doughty, C. J., & Mackey, A. (Eds.). (2021). Language aptitude: Multiple perspectives [Special issue]. *Annual Review of Applied Linguistics* (Vol. *41*).

Granena, G. (2013). Individual differences in sequence learning ability and second language acquisition in early childhood and adulthood. *Language Learning*, *63*(4), 665–703. https://doi.org/10.1111/lang.12018

Granena, G. (2014). Language aptitude and long-term achievement in early childhood L2 learners. *Applied Linguistics*, *35*(4), 483–503. https://doi.org/10.1093/applin/amu013

Granena, G. (2016). Cognitive aptitudes for implicit and explicit learning and information-processing styles: An individual differences study. *Applied PsychoLinguistics*, *37*(3), 577–600. https://doi.org/10.1017/S0142716415000120

Grigorenko, E. L., Sternberg, R. J., & Ehrman, M. E. (2000). A theory-based approach to the measurement of foreign language learning ability: The Canal-F theory and test. *Modern Language Journal*, *84*(3), 390–405. https://doi.org/10.1111/0026-7902.00076

Li, S. (2015). The associations between language aptitude and second language grammar acquisition: A meta-analytic review of five decades of research. *Applied Linguistics*, *36*(3), 385–408. https://doi.org/10.1093/applin/amu054

Li, S. (2016). The construct validity of language aptitude. *Studies in Second Language Acquisition*, *38*(4), 801–842.

Li, S. (2017). The effects of cognitive aptitudes on the process and product of L2 interaction: A synthetic review. In L. Gurzynski-Weiss (Ed.), *Expanding individual difference research in the interaction approach: Investigating learners, instructors and researchers* (pp. 41–70). John Benjamins.

Li, S. (2019). Six decades of Language aptitude research: A comprehensive and critical review. In Z. E. Wen, P. Skehan, A. Biedroń, S. Li, & R. Sparks (Eds.), *Language aptitude: Advancing theory, testing, research and practice* (pp. 78–96). Routledge.

Li, S., & DeKeyser, R. (Eds.). (2021). Implicit language aptitude: Conceptualizing the construct, validating the measures, and examining the evidence [Special issue]. *Studies in Second Language Acquisition* (Vol. *43*).

Li, S., Ellis, R., & Zhu, Y. (2019). The associations between cognitive ability and L2 development under five different instructional conditions. *Applied PsychoLinguistics*, *40*(3), 693–722. https://doi.org/10.1017/S0142716418000796

Linck, J. A., Hughes, M. M., Campbell, S. G., Silbert, N. H., Tare, M., Jackson, S. R., Smith, B. K., Bunting, M. F., & Doughty, C. J. (2013). Hi-LAB: A new measure of aptitude for high-level language proficiency. *Language Learning*, *63*(3), 530–566. https://doi.org/10.1111/lang.12011

Mitchell, R., Myles, F., & Marsden, E. (2019). *Second language learning theories*. Routledge.

Muñoz, C. (2017). Tracing trajectories of young learners: Ten years of school English learning. *Annual Review of Applied Linguistics*, *37*, 168–184. https://doi.org/10.1017/S0267190517000095

Pickering, C., & Byrne, J. (2014). The benefits of publishing systematic quantitative literature reviews for PhD candidates and other early-career researchers. *Higher Education Research & Development*, *33*(3), 534–548. https://doi.org/10.1080/07294360.2013.841651

Pimsleur, P. (1966). *Pimsleur language aptitude battery*. Harcourt, Brace & World.

Reardon, S. F., Robinson-Cimpian, J. P., & Weathers, E. S. (2014). Patterns and trends in racial/ethnic and socio-economic academic achievement gaps. In M. Ladd & E. Goertz (Eds.), *Handbook of research in education, finance and policy* (pp. 491–509). Routledge.

Robinson, P. (2012). Individual differences, aptitude complexes, SLA processes, and aptitude test development. In M. Pawlak (Ed.), *New perspectives on individual differences in language learning and teaching* (pp. 57–75). Springer.

Rogers, V. E., Meara, P., Aspinall, R., Fallon, L., Goss, T., Keey, E., & Thomas, R. (2016). Testing aptitude: Investigating Meara's (2005) LLAMA tests. *EUROSLA Yearbook*, *16*(1), 179–210. https://doi.org/10.1075/eurosla.16.07rog

Rogers, V. E., Meara, P., Barnett-Legh, T., Curry, C., & Davie, E. (2017). Examining the LLAMA aptitude tests. *Journal of the European Second Language Association*, *1*(1), 49–60. https://doi.org/10.22599/jesla.24

Saito, K. (2017). Effects of sound, vocabulary, and grammar learning aptitude on adult second language speech attainment in foreign language classrooms. *Language Learning*, *67*(3), 665–693. https://doi.org/10.1111/lang.12244

Sasaki, M. (1993a). Relationships among second language proficiency, foreign language aptitude, and intelligence: A protocol analysis. *Language Learning*, *43*(4), 469–503. https://doi.org/10.1111/j.1467-1770.1993.tb00625.x

Sasaki, M. (1993b). Relationships among second language proficiency, foreign language aptitude, and intelligence: A structural equation modeling approach. *Language Learning*, *43*(3), 313–344. https://doi.org/10.1111/j.1467-1770.1993.tb00617.x

Sasaki, M. (2012). The Modern Language Aptitude Test (paper-and-pencil version). *Language Testing*, *29*(2), 315–321. https://doi.org/10.1177/0265532211434015

Sawyer, M., & Ranta, L. (2001). Aptitude, individual differences, and instructional design. In P. Robinson (Ed.), *Cognition and second language acquisition* (pp. 319–353). Cambridge University Press.

Sheffield, S. D. (1993). Evolutionary perspectives on language: A twin study of adult foreign language learning ability [Ph.D., University of California, Berkeley]. ProQuest Dissertations and Theses. http://search.proquest.com/docview/304037971

Singleton, D. (2017). Language aptitude: Desirable trait or acquirable attribute? *Studies in Second Language Learning and Teaching*, *7*(1), 89–103. https://doi.org/10.14746/ssllt.2017.7.1.5

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford University Press.

Skehan, P. (2002). Theorising and updating aptitude. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 69–93). John Benjamins.

Sparks, R. L., & Patton, J. (2013). Relations of L1 skills and L2 aptitude to L2 anxiety on the foreign language classroom anxiety scale. *Language Learning*, *63*(4), 870–895. https://doi.org/10.1111/lang.12025

Sparks, R. L., Patton, J., Ganschow, L., & Humbach, N. (2009). Long-term crosslinguistic transfer of skills from L1 to L2. *Language Learning*, *59*(1), 203–243. https://doi.org/10.1111/j.1467-9922.2009.00504.x

Spolsky, B. (1995). Prognostication and language aptitude testing, 1925–62. *Language Testing*, *12*(3), 321–340. https://doi.org/10.1177/026553229501200304

VanPatten, B., Collopy, E., Price, J. E., Borst, S., & Qualin, A. (2013). Explicit information, grammatical sensitivity, and the first-noun principle: A cross-linguistic study in processing instruction. *The Modern Language Journal*, *97*(2), 506–527. https://doi.org/10.1111/j.1540-4781.2013.12007.x

Wen, Z. (2016). *Working memory and second language learning: Towards an integrated approach*. Multilingual Matters.

Wen, Z., Biedroń, A., & Skehan, P. (2017). Foreign language aptitude theory: Yesterday, today and tomorrow. *Language Teaching*, *50*(1), 1–31. https://doi.org/10.1017/S0261444816000276

Wen, Z., Skehan, P., Biedroń, A., Li, S., & Sparks, R. L. (Eds.). (2019). *Language aptitude: Advancing theory, testing, research and practice*. Routledge.

Winke, P. (2013). An investigation into second language aptitude for advanced Chinese language learning. *The Modern Language Journal*, *97*(1), 109–130. https://doi.org/10.1111/j.1540-4781.2013.01428.x

AUTHOR BIOGRAPHIES

James Chalmers is currently an Academic Skills Facilitator at Torrens University, Australia. He obtained his PhD in Applied Linguistics at Griffith University. His PhD thesis explored the stability of L2 aptitude and the trainability of language analytic abilities.

Susana A. Eisenchlas (PhD The University of Queensland, Australia) is currently a Senior Lecturer in Linguistics/Applied Linguistics in the School of Humanities, Languages and Social Science, Griffith University. She obtained her PhD in Linguistics at the University of Queensland. Her PhD thesis dealt with first language acquisition of Spanish syntax. Since then, she has published widely in the areas of intercultural communication, gender studies, first/second language acquisition and more recently, bilingualism and biliteracy. She is an editor of the open-access book series *Current Issues in Bilingualism* (Language Science Press), as well as an editorial board member of number of academic journals.

Andrew Munro (PhD The University of Queensland, Australia) teaches in the Spanish Studies program at Griffith University, Queensland, Australia. His research interests include Spanish language teaching and the application of rhetorical genre theory and Peircean semiotics to literary studies and cultural studies.

Andrea C. Schalley (PhD Ludwig Maximilian University of Munich, Germany) is Chair Professor in English Linguistics at Karlstad University, Sweden. She is also Co-Director of the Centre for Language and Literature Education at Karlstad University, and Research Affiliate at the Australian Research Council Centre of Excellence for the Dynamics of Language. She has published widely in high-quality outlets, on bilingualism, biliteracy, semantics and linguistic typology. Currently, she serves as editor-in-chief of the open-access book series *Current Issues in Bilingualism* (Language Science Press) and as editor of the *Journal of Home Language Research*.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

APPENDIX A

TABLE A1 Definitions of terms in SQLR coding

Category	Term	Definition
Bibliographic data	Title	Name of the source (i.e., journal article or PhD thesis).
	Author 1–3	First three listed authors of the paper in order.
	Authors additional	The remaining authors.
	Year published	The year the source was published, as reported on the paper or else as per metadata on Zotero.
	Name of publication	The name of the publication as written on the paper or else as per metadata on Zotero.
	Type of publication	Each publication was categorised as a peer reviewed journal article or a PhD dissertation.
Location	Researcher institution	The institution of each author as listed on the paper.
	Country of researcher institution	The country of the institution for each author, taken from the paper if available or from Google Maps if not.
Participants	Total number	The total number of participants whose data was collected and analysed, as reported in the source.
	Average age	The average age for the total number of participants as reported or calculated, if appropriate.
	Number of males/females	The total number of female and male participants, as reported in the paper.
	SES	The socio-economic status of participants, if reported, categorised as lower, middle, upper, or not reported. No attempt was made to standardise the measures across studies, but rather categories were reported as listed in each study.
	Education level	The highest level of education undertaken by participants at the time of the study, categorised as none, elementary, high school, university, post-graduate, or not reported. No attempt was made to standardise the measures across studies, but rather categories were listed as per the study.
	L1s	The first language of participants as reported in the study. In the case where this was not reported, the context of the study, for example, country, was used to infer the L1 and the entry was also explicitly recorded as inferred and not reported.
Data	Type of data	The type of data collected was categorised according to Dörnyei (2007):
		Quantitative = numerical data representing predetermined logical scales of values, for example, aptitude test;

(Continues)

TABLE A1 (Continued)

Category	Term	Definition
		Qualitative = verbal categories of data determined by examining the data after collection;
		Mixed = where both these types of data were included, for example, testing followed by interviews.
	Collection	The collection of data was categorised according to Dörnyei (2007):
		'longitudinal' = IF L2 aptitude testing took place before L2 learning AND L2 achievement test, especially if analysis uses L2 aptitude test scores to predict L2 achievement;
		'cross-sectional' = IF L2 aptitude testing did NOT take place before L2 learning , especially if analysis does not use L2 aptitude test scores to predict L2 achievement;
		'both' = IF L2 aptitude testing took place before L2 learning AND L2 achievement test AND analysis involves both prediction and comparison of two or more groups, where time of data collected is not important, for example, Smythe et al. (1972)
	Location of data collection	The country in which the data was collected as reported in the study. Up to four locations were recorded for one study.
	L2s	**Languages** being studied by the participants. Other options include:
		Artificial referred to studies where a modified language system was used, whether it be complete or not.
		Not reported referred to those studies that did not detail the second language being learned by participants.
		None referred to when no second language was being learned by participants.
L2 aptitude focus	L2 aptitude in study	Each study was categorised as being one of three types:
		'ID variable explaining outcome' = IF L2 aptitude is NOT the focus of the study AND L2 aptitude is a variable of interest to explain the results of the study;
		'L2 aptitude itself' = IF the study develops a new instrument for measuring L2 aptitude OR IF the study aims to expand our understanding of L2 aptitude, especially in relation to other predictors of L2 aptitude, for example, working memory, attitude/motivation, etc.;

T A B L E   A 1   (Continued)

Category	Term	Definition
		'Control variable for selecting/comparing participants' = IF L2 aptitude test scores from different treatment/participant groups in the study are analysed for statistically significant differences before the main analysis is carried out, especially if referred to as a covariate between the groups.
L2 aptitude studies	Design	The level of interaction with participants the researchers designed into the study, categorised as:
		experimental = where the study explicitly introduced a specific treatment and then compared these outcomes against a control;
		observational = where the study collected data from participants' normal course of L2 learning.
	Dependent variables	The dependent variable was determined by the statistical test used for the analysis, for example, the response variable in a general linear model. Categories covered:
		'overall L2 achievement' = IF L2 achievement is assessed as a whole without a focus on a single specific skill or area of knowledge, especially if L2 test scores are aggregated into one measure;
		'specific L2 achievement' = IF L2 achievement is assessed for a single specific skill or area of knowledge AND NOT as a whole, especially if individual sub-test scores are predicted by L2 aptitude test scores;
		'specific L2 knowledge' = IF L2 achievement was measured through accuracy scores of specific L2 structures, especially for specific grammatical structures learned through a controlled treatment;
		'specific L2 skill' = IF L2 achievement was operationalised as a specific L2 skill, for example, speaking, listening, reading (scoring could include accuracy and time measures)
		'learning ability' = IF participants had been categorised according to a diagnosis of their ability to learn, almost exclusively found in studies by Sparks, Ganschow, and colleagues;
		'L1 skills' = IF L2 aptitude test scores were used to predict standardised test scores of L1 skills and knowledge, mainly found in US studies of primary and high school children conducted by Sparks, Ganschow, and colleagues;

TABLE A1 (Continued)

Category	Term	Definition
		'L2 experience and development' = IF qualitative data that reports on the L2 learning experience and/or development of participants is used to categorise participants with L2 aptitude test scores used to explain/predict these groupings
	Control variables	Any variables used to limit the sample of participants in the study, control for differences among different groups of participants, or used as a covariant.
	Type of independent variable (IV)	The name of the explicit independent variable used in the analysis.
	Measure of L2 aptitude	The name of the specific instrument(s) used to measure L2 aptitude or L2 aptitude abilities. This includes the standardised and common tests of L2 aptitude, their adaptations (e.g. translations), or instruments explicitly designed to measure L2 aptitude or an L2 aptitude ability.

Abbreviations: SES, Socio-economic status; SQLR, systematic quantitative literature review.

APPENDIX B

B.1 REFERENCE LIST FOR PAPERS INCLUDED IN THE SQLR

Abrahamsson, N., & Hyltenstam, K. (2008). The robustness of aptitude effects in near-native second language acquisition. *Studies in Second Language Acquisition*, *30*(4), 481–509. https://doi.org/10.1017/S027226310808073X

Alderson, J. C., Clapham, C., & Steel, D. (1997). Metalinguistic knowledge, language aptitude and language proficiency. *Language Teaching Research*, *1*(2), 93–121. https://doi.org/10.1177/136216889700100202

Bain, S. K., McCallum, R. S., Bell, S. M., Cochran, J. L., & Sawyer, S. C. (2010). Foreign language learning aptitudes, attitudes, attributions, and achievement of postsecondary students identified as gifted. *Journal of Advanced Academics*, *22*(1), 130–156. https://doi.org/10.1177/1932202X1002200106

Bell, S. M., & McCallum, R. S. (2012). Do foreign language learning, cognitive, and affective variables differ as a function of exceptionality status and gender? *International Education*, *42*(1), 86–106.

Benson, S., & DeKeyser, R. (2018). Effects of written corrective feedback and language aptitude on verb tense accuracy. *Language Teaching Research*, *23*, 702–726. https://doi.org/10.1177/1362168818770921

Caruso, C. L. (1996). *Foreign language anxiety and selected learner variables in adolescent language learners* [Ph.D., The University of Wisconsin, Madison]. ProQuest Dissertations and Theses. https://search.proquest.com/docview/304293713/

Collins, B. (2018). *Sandhi-variation and the comprehension of spoken English for Japanese learners* [Ph.D., Temple University]. ProQuest Dissertations and Theses. https://search.proquest.com/docview/2046842018

Cox, J. G. (2013). *Bilingualism, aging, and instructional conditions in non-primary language development* [Ph.D., Georgetown University]. ProQuest Dissertations and Theses. https://search.proquest.com/docview/1492334739/

Cox, J. G., Lynch, J. M., Mendes, N., & Zhai, C. (2019). On bilingual aptitude for learning new languages: The roles of linguistic and nonlinguistic individual differences. *Language Learning*, *69*(2), 478–514. https://doi.org/10.1111/lang.12341

Cummins, J., & Gulutsan, M. (1975). Set, objectification and second language learning. *Journal of Psychology*, *10*(2), 91–100. https://doi.org/10.1080/00207597508247323

Curcic, M., Andringa, S., & Kuiken, F. (2019). The role of awareness and cognitive aptitudes in L2 predictive language processing. *Language Learning*, *69*(S1), 42–71. https://doi.org/10.1111/lang.12321

Dąbrowska, E. (2019). Experience, aptitude, and individual differences in linguistic attainment: A comparison of native and nonnative speakers. *Language Learning*, 69(S1), 72–100. https://doi.org/10.1111/lang.12323

Dahlen, K., & Caldwell–Harris, C. (2013). Rehearsal and aptitude in foreign vocabulary learning. *The Modern Language Journal*, 97(4), 902–916. https://doi.org/10.1111/j.1540-4781.2013.12045.x

Damiano, L. (1987). *The relationship between cognitive styles and Spanish achievement in the secondary school* [Ph.D., Fordham University]. ProQuest Dissertations and Theses. https://search.proquest.com/docview/303564705

Davies, A. (1971). Language aptitude in the first year of the U.K. secondary school. *RELC Journal*, 2(1), 4–19. https://doi.org/10.1177/003368827100200102

Demuth, K. A., & Smith, N. B. (1987). The foreign language requirement: An alternative program. *Foreign Language Annals*, 20(1), 67–77. https://doi.org/10.1111/j.1944-9720.1987.tb02919.x

Dernoshek, D. J. (1996). *The effects of a reductive grammar approach versus a traditional grammar approach on middle school Spanish I classes* [Ph.D., University of Pittsburgh]. ProQuest Dissertations and Theses. https://search.proquest.com/docview/304312457

Detschelt, F. (1993). *The relationship of learning modality, reading achievement and foreign language aptitude to selected foreign language achievement tests of second-year high school students of German* [Ph.D., University of Pittsburgh]. ProQuest Dissertations and Theses. https://search.proquest.com/docview/304097342

Doughty, C. J. (2019). Cognitive language aptitude. *Language Learning*, 69(S1), 101–126. https://doi.org/10.1111/lang.12322

Downey, D. M., Snyder, L. E., & Hill, B. (2000). College students with dyslexia: Persistent linguistic deficits and foreign language learning. *Dyslexia*, 6(2), 101–111. https://doi.org/10.1002/(SICI)1099-0909(200004/06)6:2<101::AID-DYS154>3.0.CO;2-8

Ehrman, M. E., & Oxford, R. L. (1995). Cognition plus: Correlates of language learning success. *Modern Language Journal*, 79(1), 67–89. https://doi.org/10.2307/329394

Erlam, R. (2005). Language aptitude and its relationship to instructional effectiveness in second language acquisition. *Language Teaching Research*, 9(2), 147–171. https://doi.org/10.1191/1362168805lr161oa

Faretta-Stutenberg, M., & Morgan-Short, K. (2018). The interplay of individual differences and context of learning in behavioral and neurocognitive second language development. *Second Language Research*, 34(1), 67–101. https://doi.org/10.1177/0267658316684903

Farshi, N., & Tavakoli, M. (2019). Effects of differences in language aptitude on learning grammatical collocations under elaborated input conditions. *Language Teaching Research*, 25, 476–499. https://doi.org/10.1177/1362168819858443

Feenstra, H. J. (1967). *Aptitude, attitude, and motivation in second-language acquisition* [Ph.D., The University of Western Ontario]. ProQuest Dissertations and Theses. https://search.proquest.com/docview/302299694

Gardner, R. C., Tremblay, P. F., & Masgoret, A.-M. (1997). Towards a full model of second language learning: An empirical investigation. *The Modern Language Journal*, 81(3), 344–362. https://doi.org/10.1111/j.1540-4781.1997.tb05495.x

Granena, G. (2013). Individual differences in sequence learning ability and second language acquisition in early childhood and adulthood. *Language Learning*, 63(4), 665–703. https://doi.org/10.1111/lang.12018

Granena, G. (2014). Language aptitude and long-term achievement in early childhood L2 learners. *Applied Linguistics*, 35(4), 483–503. https://doi.org/10.1093/applin/amu013

Granena, G. (2016). Cognitive aptitudes for implicit and explicit learning and information-processing styles: An individual differences study. *Applied Psycholinguistics*, 37(3), 577–600. https://doi.org/10.1017/S0142716415000120

Granena, G. (2019). Cognitive aptitudes and L2 speaking proficiency: Links between LLAMA and HI-LAB. *Studies in Second Language Acquisition*, 41(2), 313–336. https://doi.org/10.1017/S0272263118000256

Grigorenko, E. L., Sternberg, R. J., & Ehrman, M. E. (2000). A theory-based approach to the measurement of foreign language learning ability: The Canal-F theory and test. *Modern Language Journal*, 84(3), 390–405. https://doi.org/10.1111/0026-7902.00076

Horwitz, E. K. (1980). *The relationship of conceptual level to the development of communicative competence in French* [Ph.D., University of Illinois at Urbana-Champaign]. ProQuest Dissertations and Theses. https://search.proquest.com/docview/303019164

Howes, A. A. (1999). *Exploration of a tactile working memory component in foreign language learning among university students* [Psy.D., Indiana University of Pennsylvania]. ProQuest Dissertations and Theses. https://search.proquest.com/docview/304555452

Huang, B. H., Chang, Y.-H. S., Zhi, M., & Niu, L. (2018). The effect of input on bilingual adolescents' long-term language outcomes in a foreign language instruction context. *International Journal of Bilingualism*, *24*, 8–25. https://doi.org/10.1177/1367006918768311

Hughes, H. Z. (2010). *Language placement policy: Predicting strategic language learning success at the United States Air Force Academy* [Ph.D., University of Colorado at Colorado Springs]. ProQuest Dissertations and Theses. https://search.proquest.com/docview/516355911

Hwu, F., Pan, W., & Sun, S. (2014). Aptitude-treatment interaction effects on explicit rule learning: A latent growth curve analysis. *Language Teaching Research*, *18*(3), 294–319. https://doi.org/10.1177/1362168813510381

Lado, B. (2017). Aptitude and pedagogical conditions in the early development of a non-primary language. *Applied Psycholinguistics*, *38*(3), 679–701. https://doi.org/10.1017/S0142716416000394

Larson-Hall, J. (2008). Weighing the benefits of studying a foreign language at a younger starting age in a minimal input situation. *Second Language Research*, *24*(1), 35–63. https://doi.org/10.1177/0267658307082981

Li, S. (2017). An exploratory study on the role of foreign language aptitudes in instructed pragmatics learning in L2 Chinese. *Chinese as a Second Language Research*, *6*(1), 103–128. https://doi.org/10.1515/caslar-2017-0005

Li, S., Ellis, R., & Zhu, Y. (2019). The associations between cognitive ability and L2 development under five different instructional conditions. *Applied Psycholinguistics*, *40*(3), 693–722. https://doi.org/10.1017/S0142716418000796

Morgan-Short, K., Deng, Z., Brill-Schuetz, K. A., Faretta-Stutenberg, M., Wong, P. C. M., & Wong, F. C. K. (2015). A view of the neural representation of second language syntax through artificial language learning under implicit contexts of exposure. *Studies in Second Language Acquisition*, *37*(2), 383–419. https://doi.org/10.1017/S0272263115000030

Morgan-Short, K., Faretta-Stutenberg, M., Brill-Schuetz, K. A., Carpenter, H., & Wong, P. C. M. (2014). Declarative and procedural memory as individual differences in second language acquisition. *Bilingualism: Language and Cognition*, *17*(1), 56–72. https://doi.org/10.1017/S1366728912000715

Mueller, J. (2017). *An examination of the influence of age on L2 acquisition of English sound-symbolic patterns* [Ph.D., University of Maryland, College Park]. ProQuest Dissertations and Theses. https://search.proquest.com/docview/2012245985

Mueller, T. H. (1968). Programmed language instruction: Help for the linguistically "underprivileged". *Modern Language Journal*, *52*(2), 79–84. https://doi.org/10.2307/322704

Muñoz, C. (2017). Tracing trajectories of young learners: Ten years of school English learning. *Annual Review of Applied Linguistics*, *37*, 168–184. https://doi.org/10.1017/S0267190517000095

Nakata, T., & Suzuki, Y. (2018). Effects of massing and spacing on the learning of semantically related and unrelated words. *Studies in Second Language Acquisition*, *41*(2), 287–311. https://doi.org/10.1017/S0272263118000219

Nyikos, M. (1990). Sex-related differences in adult language learning: Socialization and memory factors. *Modern Language Journal*, *74*(3), 273–287. https://doi.org/10.2307/327624

O'Maley, P. J. (1993). *Second language learners in a language and culture immersion program: Longitudinal case studies in an ethnographic framework* [Ph.D., Ball State University]. ProQuest Dissertations and Theses. http://search.proquest.com/docview/304041337

Phillips, S. K. (1998). *Predictive validity for the modern language aptitude test* [Ph.D., Indiana University of Pennsylvania]. ProQuest Dissertations and Theses. http://search.proquest.com/docview/304464779

Qi, Z., Han, M., Wang, Y., de los Angeles, C., Liu, Q, Garel, K., Chen, E. S., Whitfield-Gabrieli, S., Gabrieli, J. D. E., & Perrachione, T. K. (2019). Speech processing and plasticity in the right hemisphere predict variation in adult foreign language learning. *NeuroImage*, *192*, 76–87. https://doi.org/10.1016/j.neuroimage.2019.03.008

Robinson, P. J. (1994). *Learning simple and complex second language rules under implicit, incidental, rule-search and instructed conditions* [Ph.D., University of Hawai'i at Manoa]. ProQuest Dissertations and Theses. http://search.proquest.com/docview/304107614

Robinson, P. (2005). Cognitive abilities, chunk-strength, and frequency effects in implicit artificial grammar and incidental L2 learning: Replications of Reber, Walkenfeld, and Hernstadt (1991) and Knowlton and Squire (1996) and their relevance for SLA. *Studies in Second Language Acquisition*, *27*(2), 235–268. https://doi.org/10.1017/S0272263105050126

Ruiz, S., Rebuschat, P., & Meurers, D. (2019), The effects of working memory and declarative memory on instructed second language vocabulary learning: Insights from intelligent CALL. *Language Teaching Research*, *25*, 510–539. https://doi.org/10.1177/1362168819872859

Saito, K. (2017). Effects of sound, vocabulary, and grammar learning aptitude on adult second language speech attainment in foreign language classrooms. *Language Learning*, *67*(3), 665–693. https://doi.org/10.1111/lang.12244

Saito, K. (2019). The role of aptitude in second language segmental learning: The case of Japanese learners' English /ɹ/ pronunciation attainment in classroom settings. *Applied Psycholinguistics*, *40*(1), 183–204. https://doi.org/10.1017/S0142716418000528

Saito, K., Suzukida, Y., & Sun, H. (2019). Aptitude, experience, and second language pronunciation proficiency development in classroom settings: A longitudinal study. *Studies in Second Language Acquisition*, *41*(1), 201–225. https://doi.org/10.1017/S0272263117000432

Sasaki, M. (1993). Relationships among second language proficiency, foreign language aptitude, and intelligence: A structural equation modeling approach. *Language Learning*, *43*(3), 313–344. https://doi.org/10.1111/j.1467-1770.1993.tb00617.x

Sasaki, M. (1993). Relationships among second language proficiency, foreign language aptitude, and intelligence: A protocol analysis. *Language Learning*, *43*(4), 469–503. https://doi.org/10.1111/j.1467-1770.1993.tb00625.x

Schulz, R. A. (1977). Discrete-point versus simulated communication testing in foreign languages. *Modern Language Journal*, *61*(3), 94–101. https://doi.org/10.2307/325537

Sheffield, S. D. (1993). *Evolutionary perspectives on language: A twin study of adult foreign language learning ability* [Ph.D., University of Califorfnia, Berkeley]. ProQuest Dissertations and Theses. http://search.proquest.com/docview/304037971

Sparks, R. L., Artzer, M., Ganschow, L., Siebenhar, D., Plageman, M., & Patton, J. (1998). Differences in native-language skills, foreign-language aptitude, and foreign-language grades among high-, average-, and low-proficiency foreign-language learners: Two studies. *Language Testing*, *15*(2), 181–216. https://doi.org/10.1177/026553229801500203

Sparks, R. L., & Ganschow, L. (1993). The effects of multisensory structured language instruction on native language and foreign-language aptitude skills of at-risk high-school foreign-language learners: A replication and follow-up-study. *Annals of Dyslexia*, *43*, 194–216. https://doi.org/10.1007/BF02928182

Sparks, R. L, & Ganschow, L. (1995). Parent perceptions in the screening of performance in foreign-language courses. *Foreign Language Annals*, *28*(3), 371–391. https://doi.org/10.1111/j.1944-9720.1995.tb00806.x

Sparks, R. L., & Ganschow, L. (2008). Is the foreign language classroom anxiety scale measuring anxiety or language skills? *Foreign Language Annals*, *40*(2), 260–287. https://doi.org/10.1111/j.1944-9720.2007.tb03201.x

Sparks, R. L., Ganschow, L., Artzer, M. E., Siebenhar, D., & Plageman, M. (2004). Foreign language teachers' perceptions of students' academic skills, affective characteristics, and proficiency: Replication and follow-up studies. *Foreign Language Annals*, *37*(2), 263–278. https://doi.org/10.1111/j.1944-9720.2004.tb02199.x

Sparks, R. L., Ganschow, L., Javorsky, J., Pohlman, J., & Patton, J. (1992). Identifying native language deficits in high- and low-risk foreign language learners in high school. *Foreign Language Annals*, *25*(5), 403–418. https://doi.org/10.1111/j.1944-9720.1992.tb01121.x

Sparks, R. L., Ganschow, L., Pohlman, J., Skinner, S., & Artzer, M. (1992). The effects of multisensory structured language instruction on native language and foreign language aptitude skills of at-risk high school foreign language learners. *Annals of Dyslexia*, *42*(1), 25–53. https://doi.org/10.1007/BF02654937

Sparks, R. L., Humbach, N., Patton, J., & Ganschow, L. (2011). Subcomponents of second-language aptitude and second-language proficiency. *The Modern Language Journal*, *95*(2), 253–273. https://doi.org/10.1111/j.1540-4781.2011.01176.x

Sparks, R. L., Javorsky, J., Ganschow, L., Pohlman, J., & Patton, J. (1992). Test comparisons among students identified as high-risk, low-risk, and learning disabled in high school foreign language courses. *The Modern Language Journal*, *76*(2), 142–159. https://doi.org/10.2307/329768

Sparks, R. L., Luebbers, J., Castañeda, M., & Patton, J. (2018). High school Spanish students and foreign language reading anxiety: Déjà vu all over again all over again. *The Modern Language Journal*, *102*(3), 533–556. https://doi.org/10.1111/modl.12504

Sparks, R. L., & Patton, J. (2013). Relationship of L1 skills and L2 aptitude to L2 anxiety on the foreign language classroom anxiety scale. *Language Learning*, *63*(4), 870–895. https://doi.org/10.1111/lang.12025

Sparks, R. L., Patton, J., Ganschow, L., & Humbach, N. (2009). Long-term crosslinguistic transfer of skills from L1 to L2. *Language Learning*, *59*(1), 203–243. https://doi.org/10.1111/j.1467-9922.2009.00504.x

Sparks, R. L., Patton, J., Ganschow, L., & Humbach, N. (2009). Long-term relationships among early first language skills, second language aptitude, second language affect, and later second language proficiency. *Applied Psycholinguistics*, *30*(4), 725–755. https://doi.org/10.1017/S0142716409990099

Sparks, R. L., Patton, J., & Luebbers, J. (2018). L2 anxiety and the foreign language reading anxiety scale: Listening to the evidence. *Foreign Language Annals*, *51*(4), 738–762. https://doi.org/10.1111/flan.12361

Sparks, R. L., Patton, J., & Luebbers, J. (2019). Individual differences in L2 achievement mirror individual differences in L1 skills and L2 aptitude: Crosslinguistic transfer of L1 to L2 skills. *Foreign Language Annals*, *52*(2), 255–283. https://doi.org/10.1111/flan.12390

Sparks, R. L., Philips, L., Ganschow, L., & Javorsky, J. (1999). Students classified as LD and the college foreign language requirement: A quantitative analysis. *Journal of Learning Disabilities*, *32*(6), 566–580. https://doi.org/10.1177/002221949903200608

Sparks, R. L., Philips, L. G., & Javorsky, J. (2002). Students classified as LD who received course substitutions for the college foreign language requirement: A replication study. *Journal of Learning Disabilities*, *35*(6), 482–500. https://doi.org/10.1177/00222194020350060101

Stefanou, C., & Révész, A. (2015). Direct written corrective feedback, learner differences, and the acquisition of second language article use for generic and specific plural reference. *Modern Language Journal*, *99*(2), 263–282. https://doi.org/10.1111/modl.12212

Suzuki, Y., & DeKeyser, R. (2017a). Exploratory research on second language practice distribution: An aptitude × treatment interaction. *Applied Psycholinguistics*, *38*(1), 27–56. https://doi.org/10.1017/S0142716416000084

Suzuki, Y., & DeKeyser, R. (2017b). The interface of explicit and implicit knowledge in a second language: Insights from individual differences in cognitive aptitudes. *Language Learning*, *67*(4), 747–790. https://doi.org/10.1111/lang.12241

Takeuchi-Furuya, R. (1993). *A study of the naturalistic acquisition of Japanese by two native speakers of English in Japan* [Ed.D., Temple University]. ProQuest Dissertations and Theses. http://search.proquest.com/docview/304077214

Theivananthampillai, K. T., & Baba, T. L. (1984). Student variables in second language learning. *RELC Journal*, *15*(1), 50–63. https://doi.org/10.1177/003368828401500103

Thompson, A. S. (2013). The interface of language aptitude and multilingualism: Reconsidering the bilingual/multilingual dichotomy. *The Modern Language Journal*, *97*(3), 685–701. https://doi.org/10.1111/j.1540-4781.2013.12034.x

Tolentino, L. C., & Tokowicz, N. (2014). Cross-language similarity modulates effectiveness of second language grammar instruction. *Language Learning*, *64*(2), 279–309. https://doi.org/10.1111/lang.12048

Vandergrift, L., & Baker, S. (2015). Learner variables in second language listening comprehension: An exploratory path analysis. *Language Learning*, *65*(2), 390–416. https://doi.org/10.1111/lang.12105

VanPatten, B., & Borst, S. (2012). The roles of explicit information and grammatical sensitivity in processing instruction: Nominative-accusative case marking and word order in German L2. *Foreign Language Annals*, *45*(1), 92–109. https://doi.org/10.1111/j.1944-9720.2012.01169.x

VanPatten, B., Collopy, E., Price, J. E., Borst, S., & Qualin, A. (2013). Explicit information, grammatical sensitivity, and the first-noun principle: A cross-linguistic study in processing instruction. *The Modern Language Journal*, *97*(2), 506–527. https://doi.org/10.1111/j.1540-4781.2013.12007.x

Winke, P. (2013). An investigation into second language aptitude for advanced Chinese language learning. *The Modern Language Journal*, *97*(1), 109–130. https://doi.org/10.1111/j.1540-4781.2013.01428.x

Wistner, B. (2014). *Effects of metalinguistic knowledge and language aptitude on second language learning* [Ph.D., Temple University]. ProQuest Dissertations and Theses. https://search.proquest.com/docview/1622286083

Yalçin, Ş., & Spada, N. (2016). Language aptitude and grammatical difficulty. *Studies in Second Language Acquisition*, *38*(S2), 239–263. https://doi.org/10.1017/S0272263115000509

Yi, W. (2018). Statistical sensitivity, cognitive aptitudes, and processing of collocations. *Studies in Second Language Acquisition*, *40*(4), 831–856. https://doi.org/10.1017/S0272263118000141

Yilmaz, Y., & Granena, G. (2016). The role of cognitive aptitudes for explicit language learning in the relative effects of explicit and implicit feedback. *Bilingualism: Language and Cognition*, *19*(1), 147–161. https://doi.org/10.1017/S136672891400090X

Yilmaz, Y., & Granena, G. (2019). Cognitive individual differences as predictors of improvement and awareness under implicit and explicit feedback conditions. *Modern Language Journal*, *103*(3), 686–702. https://doi.org/10.1111/modl.12587